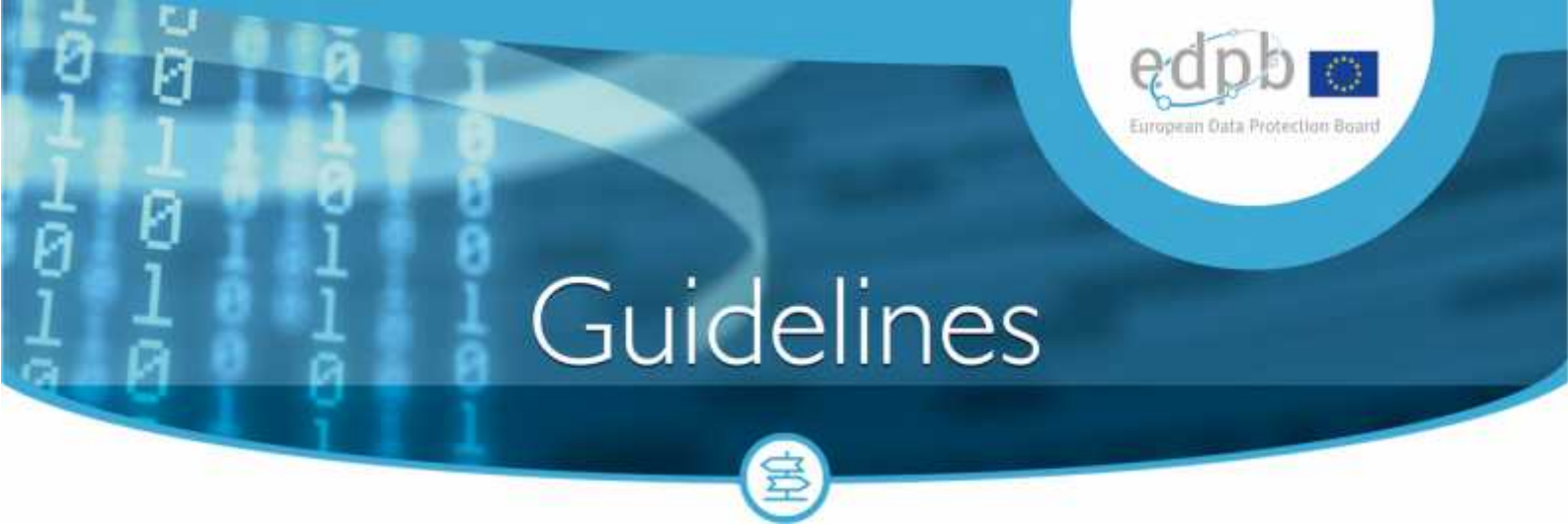

Guidelines 4/2019 on Article 25

Data Protection by Design and by Default

Version 2.0

Adopted on 20 October 2020

Version history

Version 1.0	13 November 2019	Adoption of the Guidelines for public consultation
Version 2.0	20 October 2020	Adoption of the Guidelines by the EDPB after public consultation

Table of contents

The European Data Protection Board

Having regard to Article 70 (1e) of the Regulation 2016/679/EU of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC], (hereinafter "GDPR"),

Having regard to the EEA Agreement and in particular to Annex XI and Protocol 37 thereof, as amended by the Decision of the EEA joint Committee No 154/2018 of 6 July 2018,

Having regard to Article 12 and Article 22 of its Rules of Procedure,

HAS ADOPTED THE FOLLOWING GUIDELINES

Executive summary

In an increasingly digital world, adherence to Data Protection by Design and by Default requirements plays a crucial part in promoting privacy and data protection in society. It is therefore essential that controllers take this responsibility seriously and implement the GDPR obligations when designing processing operations.

These Guidelines give general guidance on the obligation of Data Protection by Design and by Default (henceforth "DPbDD") set forth in Article 25 in the GDPR. DPbDD is an obligation for all controllers, irrespective of size and varying complexity of processing. To be able to implement the requirements of DPbDD, it is crucial that the controller understands the data protection principles and the data subject's rights and freedoms.

The core obligation is the implementation of *appropriate* measures and necessary safeguards that provide *effective implementation* of the *data protection principles* and*,* consequentially*, data subjects' rights and freedoms by design and by default*. Article 25 prescribes both design and default elements that should be taken into account. Those elements, will be further elaborated in these Guidelines.

Article 25(1) stipulates that controllers should consider DPbDD early on when they plan a new processing operation. Controllers shall implement DPbDD *before* processing, and also *continually* at the time of processing, by regularly reviewing the effectiveness of the chosen measures and safeguards. DPbDD also applies to existing systems that are processing personal data.

The Guidelines also contain guidance on how to effectively implement the data protection principles in Article 5, listing key design and default elements as well as practical cases for illustration. The controller should consider the appropriateness of the suggested measures in the context of the particular processing in question.

The EDPB provides recommendations on how controllers, processors and producers can cooperate to achieve DPbDD. It encourages the controllers in industry, processors, and producers to use DPbDD as a means to achieve a competitive advantage when marketing their products towards controllers and data subjects. It also encourages all controllers to make use of certifications and codes of conduct.

1 SCOPE

1. The Guidelines focus on controllers' implementation of DPbDD based on the obligation in Article 25 of the GDPR.[1] Other actors, such as processors and producers of products, services and applications (henceforth "producers"), who are not directly addressed in Article 25, may also find these Guidelines useful in creating GDPR compliant products and services that enable controllers to fulfil their data protection obligations.[2] Recital 78 of the GDPR adds that DPbDD should be taken into consideration in the context of public tenders. Despite all controllers having the duty to integrate DPbDD into their processing activities, this provision fosters the adoption of the data protection principles, where public administrations should lead by example. The controller is responsible for the fulfilment of the DPbDD obligations for the processing carried out by their processors and sub-processors, they should therefore take this into account when contracting with these parties.

2. The requirement described in Article 25 is for controllers to have data protection designed into the processing of personal data and as a default setting and this applies throughout the processing lifecycle. DPbDD is also a requirement for processing systems pre-existing before the GDPR entered into force. Controllers must have the processing consistently updated in line with the GDPR. For more information on how to maintain an existing system in line with DPbDD, see subchapter 2.1.4 of these Guidelines. The core of the provision is to ensure *appropriate* and *effective* data protection both by *design* and by *default,* which means that controllers should be able to demonstrate that they have the appropriate measures and safeguards in the processing to ensure that the data protection principles and the rights and freedoms of data subjects are effective.

3. Chapter 2 of the Guidelines focuses on an interpretation of the requirements set forth by Article 25 and explores the legal obligations introduced by the provision. Examples on how to apply DPbDD in the context of specific data protection principles are provided in Chapter 3.

4. The Guidelines address the possibility to establish a certification mechanism to demonstrate compliance with Article 25 in Chapter 4, as well as how the Article may be enforced by supervisory authorities in Chapter 5. Finally, the Guidelines provide stakeholders with further recommendations on how to successfully implement DPbDD. The EDPB recognizes the challenges for small and medium enterprises (henceforth "SMEs") to fully comply with the obligations of DPbDD, and provides additional recommendations specifically to SMEs in Chapter 6.

2 ANALYSIS OF ARTICLE 25(1) AND (2) DATA PROTECTION BY DESIGN AND BY DEFAULT

5. The aim of this Chapter is to explore and provide guidance on the requirements to data protection by design in Article 25(1) and to data protection by default in Article 25(2) respectively. Data protection

[1] The interpretations provided herein equally apply to Article 20 of Directive (EU) 2016/680, and Article 27 of Regulation 2018/1725.

[2] Recital 78 GDPR clearly states this need: "*When developing, designing, selecting and using applications, services and products that are based on the processing of personal data or process personal data to fulfil their task, producers of the products, services and applications should be encouraged to take into account the right to data protection when developing and designing such products, services and applications and, with due regard to the "state of the art", to make sure that controllers and processors are able to fulfil their data protection obligations*".

by design and data protection by default are complementary concepts, which mutually reinforce each other. Data subjects will benefit more from data protection by default if data protection by design is concurrently implemented – and vice versa.

6. DPbDD is a requirement for all controllers, including small businesses and multinational companies alike. That being the case, the complexity of implementing DPbDD may vary based on the individual processing operation. Regardless of the size however, in all cases, positive benefits for controller and data subject can be achieved by implementing DPbDD.

2.1 Article 25(1): Data protection by design

2.1.1 Controller's obligation to implement appropriate technical and organisational measures and necessary safeguards into the processing

7. In line with Article 25(1) the controller shall implement *appropriate* technical and organisational *measures* which are designed to implement the data protection principles and to integrate the *necessary safeguards* into the processing in order to meet the requirements and protect the rights and freedoms of data subjects. Both appropriate measures and necessary safeguards are meant to serve the same purpose of protecting the rights of data subjects and ensuring that the protection of their personal data is built into the processing.

8. *Technical and organizational measures* and necessary *safeguards* can be understood in a broad sense as any method or means that a controller may employ in the processing. Being *appropriate* means that the measures and necessary safeguards should be suited to achieve the intended purpose, i.e. they must implement the data protection principles *effectively*[3]. The requirement to appropriateness is thus closely related to the requirement of effectiveness.

9. A technical or organisational measure and safeguard can be anything from the use of advanced technical solutions to the basic training of personnel. Examples that may be suitable, depending on the context and risks associated with the processing in question, includes pseudonymization of personal data[4]; storing personal data available in a structured, commonly machine readable format; enabling data subjects to intervene in the processing; providing information about the storage of personal data; having malware detection systems; training employees about basic "cyber hygiene"; establishing privacy and information security management systems, obligating processors contractually to implement specific data minimisation practices, etc.

10. Standards, best practices and codes of conduct that are recognized by associations and other bodies representing categories of controllers can be helpful in determining appropriate measures. However, the controller must verify the appropriateness of the measures for the particular processing in question.

2.1.2 Designed to implement the data protection principles in an effective manner and protecting data subjects' rights and freedoms

11. The *data protection principles* are in Article 5 (henceforth "the principles"), the *data subjects' rights and freedoms* are the fundamental rights and freedoms of natural persons, and in particular their right to the protection of personal data, whose protection is named in Article 1(2) as the objective of the

[3] "Effectiveness" is addressed below in subchapter 2.1.2.
[4] Defined in Article 4(5) GDPR.

GDPR (henceforth "the rights")[5]. Their precise formulation can be found in the EU Charter of Fundamental Rights. It is essential for the controller to have an understanding of the meaning of *the principles* and *the rights* as the basis for the protection offered by the GDPR, specifically by the DPbDD obligation.

12. When implementing the appropriate technical and organisational measures, it is with respect to the effective implementation of each of the aforementioned principles and the ensuing protection of rights that the measures and safeguards should be *designed*.

Addressing effectiveness

13. Effectiveness is at the heart of the concept of data protection by design. The requirement to implement the principles in an effective manner means that controllers must implement the necessary measures and safeguards to protect these principles, in order to secure the rights of data subjects. Each implemented measure should produce the intended results for the processing foreseen by the controller. This observation has two consequences.

14. First, it means that Article 25 does not require the implementation of any specific technical and organizational measures, rather that the chosen measures and safeguards should be specific to the implementation of data protection principles into the particular processing in question. In doing so, the measures and safeguards should be designed to be robust and the controller should be able to implement further measures in order to scale to any increase in risk[6]. Whether or not measures are effective will therefore depend on the context of the processing in question and an assessment of certain elements that should be taken into account when determining the means of processing. The aforementioned elements will be addressed below in subchapter 2.1.3.

15. Second, controllers should be able to demonstrate that the principles have been maintained.

16. The implemented measures and safeguards should achieve the desired effect in terms of data protection, and the controller should have documentation of the implemented technical and organizational measures.[7] To do so, the controller may determine appropriate key performance indicators (KPI) to demonstrate the effectiveness. A KPI is a measurable value chosen by the controller that demonstrates how effectively the controller achieves their data protection objective. KPIs may be *quantitative*, such as the percentage of false positives or false negatives, reduction of complaints, reduction of response time when data subjects exercise their rights; or *qualitative*, such as evaluations of performance, use of grading scales, or expert assessments. Alternatively to KPIs, controllers may be able to demonstrate the effective implementation of the principles by providing the rationale behind their assessment of the effectiveness of the chosen measures and safeguards.

2.1.3 Elements to take into account

17. Article 25 (1) lists elements that the controller has to take into account when determining the measures of a specific processing operation. In the following, we will provide guidance on how to apply

[5] See Recital 4 of the GDPR.

[6] "Fundamental principles applicable to the controllers (i.e. legitimacy, data minimisation, purpose limitation, transparency, data integrity, data accuracy) should remain the same, whatever the processing and the risks for the data subjects. However, due regard to the nature and scope of such processing have always been an integral part of the application of those principles, so that they are inherently scalable." Article 29 Working Party. "Statement on the role of a risk-based approach in data protection legal frameworks". WP 218, 30 May 2014, p. 3. ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp218_en.pdf

[7] See Recitals 74 and 78.

these elements in the design process, which includes design of the default settings. These elements all contribute to determine whether a measure is appropriate to effectively implement the principles. Thus, each of these elements is not a goal in and of themselves, but are factors to be considered together to reach the objective.

2.1.3.1 "state of the art"

18. The concept of "state of the art" is present in various EU acquis, e.g. environmental protection and product safety. In the GDPR, reference to the "state of the art"[8] is made not only in Article 32, for security measures,[9][10] but also in Article 25, thus extending this benchmark to all technical and organisational measures embedded in the processing.

19. In the context of Article 25, the reference to "state of the art" imposes an obligation on controllers, when determining the appropriate technical and organisational measures, **to take account of the current progress in technology** that is available in the market. The requirement is for controllers to have knowledge of, and stay up to date on technological advances; how technology can present data protection risks or opportunities to the processing operation; and how to implement and update the measures and safeguards that *secure effective implementation* of the principles and rights of data subjects taking into account the evolving technological landscape.

20. The "state of the art" is a dynamic concept that cannot be statically defined at a fixed point in time, but should be assessed *continuously* in the context of technological progress. In the face of technological advancements, a controller could find that a measure that once provided an adequate level of protection no longer does. Neglecting to keep up to date with technological changes could therefore result in a lack of compliance with Article 25.

21. The "state of the art" criterion does not only apply to technological measures, but also to organisational ones. Lack of appropriate organisational measures can lower or even completely undermine the effectiveness of a chosen technology. Examples of organisational measures can be adoption of internal policies; up-to date training on technology, security and data protection; and IT security governance and management policies.

22. Existing and recognized frameworks, standards, certifications, codes of conduct, etc. in different fields may play a role in indicating the current "state of the art" within the given field of use. Where such standards exist and provide a high level of protection for the data subject in compliance with – or go beyond – legal requirements, controllers should take them into account in the design and implementation of data protection measures.

2.1.3.2 "cost of implementation"

23. The controller may take the cost of implementation into account when choosing and applying appropriate technical and organisational measures and necessary safeguards that effectively

[8] See German Federal Constitutional Court's "Kalkar" decision in 1978: https://germanlawarchive.iuscomp.org/?p=67 may provide the foundation for a methodology for an objective definition of the concept. On that basis, the "state of the art" technology level would be identified between the "existing scientific knowledge and research" technology level and the more established "generally accepted rules of technology". The "state of the art" can hence be identified as the technology level of a service or technology or product that exists in the market and is most effective in achieving the objectives identified.
[9] https://www.enisa.europa.eu/news/enisa-news/what-is-state-of-the-art-in-it-security
[10] www.teletrust.de/en/publikationen/broschueren/state-of-the-art-in-it-security/

implement the principles in order to protect the rights of data subjects. The cost refers to resources in general, including time and human resources.

24. The cost element does not require the controller to spend a disproportionate amount of resources when alternative, less resource demanding, yet effective measures exist. However, the cost of implementation is a factor to be considered to implement data protection by design rather than a ground to not implement it.

25. Thus, the chosen measures shall ensure that the processing activity foreseen by the controller does not process personal data in violation of the principles, independent of cost. Controllers should be able to manage the overall costs to be able to effectively implement all of the principles and, consequentially, protect the rights.

2.1.3.3 "nature, scope, context and purpose of processing"

26. Controllers must take into consideration the nature, scope, context and purpose of processing when determining needed measures.

27. These factors should be interpreted consistently with their role in other provisions of the GDPR, such as Articles 24, 32 and 35, with the aim of designing data protection principles into the processing.

28. In short, the concept of **nature** can be understood as the inherent[11] characteristics of the processing. The **scope** refers to the size and range of the processing. The **context** relates to the circumstances of the processing, which may influence the expectations of the data subject, while the **purpose** pertains to the aims of the processing.

2.1.3.4 "risks of varying likelihood and severity for rights and freedoms of natural persons posed by the processing"

29. The GDPR adopts a coherent risk based approach in many of its provisions, in Articles 24, 25, 32 and 35, with a view to identifying appropriate technical and organisational measures to protect individuals, their personal data and complying with the requirements of the GDPR. The assets to protect are always the same (the individuals, via the protection of their personal data), against the same risks (to individuals' rights), taking into account the same conditions (nature, scope, context and purposes of processing).

30. When performing the risk analysis for compliance with Articles 25, the controller has to identify the risks to the rights of data subjects that a violation of the principles presents, and determine their likelihood and severity in order to implement measures to effectively mitigate the identified risks. A systematic and thorough evaluation of the processing is crucial when doing risk assessments. For example, a controller assesses the particular risks associated with a lack of freely given consent, which constitutes a violation of the lawfulness principle, in the course of the processing of personal data of children and young people under 18 as a vulnerable group, in a case where no other legal ground exists, and implements appropriate measures to address and effectively mitigate the identified risks associated with this group of data subjects.

[11] Examples are special categories personal data, automatic decision-making, skewed power relations, unpredictable processing, difficulties for the data subject to exercise the rights, etc.

31. The "EDPB Guidelines on Data Protection Impact Assessment (DPIA)",[12] which focus on determining whether a processing operation is likely to result in a high risk to the data subject or not, also provide guidance on how to assess data protection risks and how to carry out a data protection risk assessment. These Guidelines may also be useful during the risk assessment in all the articles mentioned above, including Article 25.

32. The risk based approach does not exclude the use of baselines, best practices and standards. These might provide a useful toolbox for controllers to tackle similar risks in similar situations (nature, scope, context and purpose of processing). Nevertheless, the obligation in Article 25 (as well as Articles 24, 32 and 35(7)(c)) to take into account "*risks of varying likelihood and severity for rights and freedoms of natural persons posed by the processing*" remains. Therefore, controllers, although supported by such tools, must always carry out a data protection risk assessment on a case by case basis for the processing activity at hand and verify the effectiveness of the appropriate measures and safeguards proposed. A DPIA, or an update to an existing DPIA, may then additionally be required.

2.1.4 Time aspect

2.1.4.1 At the time of the determination of the means for processing

33. Data protection by design shall be implemented "*at the time of determination of the means for processing*".

34. The "*means for processing*" range from the general to the detailed design elements of the processing, including the architecture, procedures, protocols, layout and appearance.

35. The "*time of determination of the means for processing*" refers to the period of time when the controller is deciding how the processing will be conducted and the manner in which the processing will occur and the mechanisms which will be used to conduct such processing. It's in the process of making such decisions that the controller must assess the appropriate measures and safeguards to effectively implement the principles and rights of data subjects into the processing, and take into account elements such as the state of the art, cost of implementation, nature, scope, context and purpose, and risks. This includes the time of procuring and implementing data processing software, hardware, and services.

36. Early consideration of DPbDD is crucial for a successful implementation of the principles and protection of the rights of the data subjects. Moreover, from a cost-benefit perspective, it is also in controllers' interest to take DPbDD into account sooner rather than later, as it could be challenging and costly to make later changes to plans that have already been made and processing operations that have already been designed.

2.1.4.2 At the time of the processing itself (maintenance and review of data protection requirements)

37. Once the processing has started the controller has a continued obligation to maintain DPbDD, i.e. the continued effective implementation of the principles in order to protect the rights, staying up to date on the state of the art, reassessing the level of risk, etc. The nature, scope and context of processing operations, as well as the risk may change over the course of processing, which means that the

[12] Article 29 Working Party "Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679". WP 248 rev.01, 4 October 2017. ec.europa.eu/newsroom/document.cfm?doc_id=47711 - endorsed by the EDPB.

controller must re-evaluate their processing operations through regular reviews and assessments of the effectiveness of their chosen measures and safeguards.

38. The obligation to maintain, review and update, as necessary, the processing operation also applies to pre-existing systems. This means that legacy systems designed before the GDPR entered into force are required to undergo reviews and maintenance to ensure the implementation of measures and safeguards that implement the principles and rights of data subjects in an effective manner, as outlined in these Guidelines.

39. This obligation also extends to any processing carried out by means of data processors. Processors' operations should be regularly reviewed and assessed by the controllers to ensure that they enable continuous compliance with the principles and allow the data controller to fulfil its obligations in this respect.

2.2 Article 25(2): Data protection by default

2.2.1 By default, only personal data which are necessary for each specific purpose of the processing are processed

40. A "default", as commonly defined in computer science, refers to the pre-existing or preselected value of a configurable setting that is assigned to a software application, computer program or device. Such settings are also called "presets" or "factory presets", especially for electronic devices.

41. Hence, the term "by default" when processing personal data, refers to making choices regarding configuration values or processing options that are set or prescribed in a processing system, such as a software application, service or device, or a manual processing procedure that affect the amount of personal data collected, the extent of their processing, the period of their storage and their accessibility.

42. The controller should choose and be accountable for implementing default processing settings and options in a way that only processing that is strictly necessary to achieve the set, lawful purpose is carried out by default. Here, controllers should rely on their assessment of the necessity of the processing with regards to the legal grounds of Article 6(1). This means that by default, the controller shall not collect more data than is necessary, they shall not process the data collected more than is necessary for their purposes, nor shall they store the data for longer than necessary. The basic requirement is that data protection is built into the processing by default.

43. The controller is required to predetermine for which specified, explicit and legitimate purposes the personal data is collected and processed.[13] The measures must by default be appropriate to ensure that only personal data which are necessary for each specific purpose of processing are being processed. The EDPS "Guidelines to assess necessity and proportionality of measures that limit the

[13] Art. 5(1)(b), (c), (d), (e) GDPR.

right to data protection of personal data" can be useful also to decide which data is necessary to process in order to achieve a specific purpose.[14] [15] [16]

44. If the controller uses third party software or off-the-shelf software, the controller should carry out a risk assessment of the product and make sure that functions that do not have a legal basis or are not compatible with the intended purposes of processing are switched off.

45. The same considerations apply to organisational measures supporting processing operations. They should be designed to process, at the outset, only the minimum amount of personal data necessary for the specific operations. This should be particularly considered when allocating data access to staff with different roles and different access needs.

46. Appropriate "technical and organisational measures" in the context of data protection by default is thus understood in the same way as discussed above in subchapter 2.1.1, but applied specifically to implementing the principle of data minimisation.

47. The aforementioned obligation to only process personal data which are necessary for each specific purpose applies to the following elements.

2.2.2 Dimensions of the data minimisation obligation

48. Article 25 (2) lists the dimensions of the data minimisation obligation for default processing, by stating that the obligation applies to the amount of personal data collected, the extent of their processing, the period of their storage and their accessibility.

2.2.2.1 "amount of personal data collected"

49. Controllers should consider both the volume of personal data, as well as the types, categories and level of detail of personal data required for the processing purposes. Their design choices should take into account the increased risks to the principles of integrity and confidentiality, data minimisation and storage limitation when collecting large amounts of detailed personal data, and compare it to the reduction in risks when collecting smaller amounts and/or less detailed information about data subjects. In any case, the default setting shall not include collection of personal data that is not necessary for the specific processing purpose. In other words, if certain categories of personal data are unnecessary or if detailed data isn't needed because less granular data is sufficient, then any surplus personal data shall not be collected.

50. The same default requirements apply to services independent of what platform or device in use, only the necessary personal data for the given purpose can be collected.

[14] EDPS. "Guidelines on assessing the necessity and proportionality of measures that limit the right to data protection". 25 February 2019. edps.europa.eu/sites/edp/files/publication/19-02-25_proportionality_guidelines_en.pdf

[15] See also EDPS. "Assessing the necessity of measures that limit the fundamental right to the protection of personal data: A Toolkit" https://edps.europa.eu/data-protection/our-work/publications/papers/necessity-toolkit_en

[16] For more information on necessity, see Article 29 Working Party. "Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC". WP 217, 9 April 2014. ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp217_en.pdf

2.2.2.2 *"the extent of their processing"*

51. Processing[17] operations performed on personal data shall be limited to what is necessary. Many processing operations may contribute to a processing purpose. Nevertheless, the fact that certain personal data is necessary to fulfil a purpose does not mean that all types of, and frequencies of, processing operations may be carried out on the data. Controllers should also be careful not to extend the boundaries of "compatible purposes" of Article 6(4), and have in mind what processing will be within the reasonable expectations of data subjects.

2.2.2.3 *"the period of their storage"*

52. Personal data collected shall not be stored if it is not necessary for the purpose of the processing and there is no other compatible purpose and legal ground according to Article 6(4). Any retention should be objectively justifiable as necessary by the data controller in accordance with the accountability principle.

53. The controller shall limit the retention period to what is necessary for the purpose. If personal data is no longer necessary for the purpose of the processing, then it shall by default be deleted or anonymized. The length of the period of retention will therefore depend on the purpose of the processing in question. This obligation is directly related to the principle of storage limitation in Article 5(1)(e), and shall be implemented by default, i.e. the controller should have systematic procedures for data deletion or anonymization embedded in the processing.

54. Anonymization[18] of personal data is an alternative to deletion, provided that all the relevant contextual elements are taken into account and the likelihood and severity of the risk, including the risk of re-identification, are regularly assessed.[19]

2.2.2.4 *"their accessibility"*

55. The controller should limit who has access and which types of access to personal data based on an assessment of necessity, and also make sure that personal data is in fact accessible to those who need it when necessary, for example in critical situations. Access controls should be observed for the whole data flow during the processing.

56. Article 25(2) further states that personal data shall not be made accessible, without the individual's intervention, to an indefinite number of natural persons. The controller shall by default limit accessibility and give the data subject the possibility to intervene before publishing or otherwise making available personal data about the data subject to an indefinite number of natural persons.

57. Making personal data available to an indefinite number of persons may result in even further dissemination of the data than initially intended. This is particularly relevant in the context of the Internet and search engines. This means that controllers should by default give data subjects an

[17] According to Art. 4(2) GDPR, this includes collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction.
[18] Article 29 Working Party. "Opinion 05/2014 on Anonymisation Techniques". WP 216, 10 April 2014. ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf
[19] Please see Art. 4(1) GDPR, Recital 26 GDPR, Article 29 Working Party "Opinion 05/2014 on Anonymisation Techniques". Please also see the subsection on "storage limitation" in section 3 of this document, referring to the need for the controller to ensure the effectiveness of the implemented anonymisation technique(s).

opportunity to intervene before personal data is made available on the open Internet. This is particularly important when it comes to children and vulnerable groups.

58. Depending on the legal grounds for processing, the opportunity to intervene could vary based on the context of the processing. For example, to ask for consent to make the personal data publicly accessible, or to have privacy settings so that data subjects themselves can control public access.

59. Even in the event that personal data is made available publicly with the permission and understanding of a data subject, it does not mean that any other controller with access to the personal data may freely process it themselves for their own purposes – they must have their own legal basis.[20]

3 IMPLEMENTING DATA PROTECTION PRINCIPLES IN THE PROCESSING OF PERSONAL DATA USING DATA PROTECTION BY DESIGN AND BY DEFAULT

60. In all stages of design of the processing activities, including procurement, tenders, outsourcing, development, support, maintenance, testing, storage, deletion, etc., the controller should take into account and consider the various elements of DPbDD which will be illustrated by examples in this chapter in the context of implementation of the principles.[21] [22] [23]

61. Controllers need to implement the principles to achieve DPbDD. These principles include: transparency, lawfulness, fairness, purpose limitation, data minimisation, accuracy, storage limitation, integrity and confidentiality, and accountability. These principles are outlined in Article 5 and Recital 39 of the GDPR. To have a complete understanding of how to implement DPbDD, the importance of understanding the meaning of each of the principles is emphasised.

62. When presenting examples of how to operationalize DPbDD we have made lists of **key DPbDD elements** for each of the principles. The examples, while highlighting the specific data protection principle in question, may overlap with other closely related principles as well. The EDPB underlines that the key elements and the examples presented hereunder are neither exhaustive nor binding, but are meant as guiding elements for each of the principles. Controllers need to assess how to guarantee compliance with the principles in the context of the concrete processing operation in question.

63. While this section focuses on the implementation of the principles, the controller should also implement *appropriate* and *effective* ways to protect data subjects' rights, also according to Chapter III in the GDPR where this is not already mandated by the principles themselves.

64. The accountability principle is overarching: it requires the controller to be responsible choosing the necessary technical and organisational measures.

[20] See Case of Satakunnan Markkinapörssi Oy and Satamedia Oy v. Finland no. 931/13.
[21] More examples can be found in Norwegian Data Protection Authority. "Software Development with Data Protection by Design and by Default". 28 November 2017. www.datatilsynet.no/en/about-privacy/virksomhetenes-plikter/innebygd-personvern/data-protection-by-design-and-by-default/?id=7729
[22] https://www.cnil.fr/en/cnil-publishes-gdpr-guide-developers
[23] https://www.aepd.es/sites/default/files/2019-12/guia-privacidad-desde-diseno_en.pdf

3.1 Transparency[24]

65. The controller must be clear and open with the data subject about how they will collect, use and share personal data. Transparency is about enabling data subjects to understand, and if necessary, make use of their rights in Articles 15 to 22. The principle is embedded in Articles 12, 13, 14 and 34. Measures and safeguards put in place to support the principle of transparency should also support the implementation of these Articles.

66. Key design and default elements for the principle of transparency may include:

- Clarity – Information shall be in clear and plain language, concise and intelligible.
- Semantics – Communication should have a clear meaning to the audience in question.
- Accessibility - Information shall be easily accessible for the data subject.
- Contextual – Information should be provided at the relevant time and in the appropriate form.
- Relevance – Information should be relevant and applicable to the specific data subject.
- Universal design – Information shall be accessible to all data subjects, include use of machine readable languages to facilitate and automate readability and clarity.
- Comprehensible – Data subjects should have a fair understanding of what they can expect with regards to the processing of their personal data, particularly when the data subjects are children or other vulnerable groups.
- Multi-channel – Information should be provided in different channels and media, not only the textual, to increase the probability for the information to effectively reach the data subject.
- Layered – The information should be layered in a manner that resolves the tension between completeness and understanding, while accounting for data subjects' reasonable expectations.

Example[25]

A controller is designing a privacy policy on their website in order to comply with the requirements of transparency. The privacy policy should not contain a lengthy bulk of information that is difficult for the average data subject to penetrate and understand. It shall be written in clear and concise language and make it easy for the user of the website to understand how their personal data is processed. The controller therefore provides information in a layered manner, where the most important points are highlighted. More detailed information is made easily available. Drop-down menus and links to other pages are provided to further explain the various items, and concepts used in the policy. The controller also makes sure that the information is provided in a multi-channel manner, providing video clips to explain the most important points of the written information. Synergy between the various pages is vital to ensure that the layered approach does not heighten confusion, rather reduce it.

The privacy policy should not be difficult for data subjects to access. The privacy policy is thus made available and visible on all web-pages of the site in question, so that the data subject is always only one click away from accessing the information. The information provided is also designed in accordance with the best practices and standards of universal design to make it accessible to all.

[24] Elaboration on how to understand the concept of transparency can be found in Article 29 Working Party. "Guidelines on transparency under Regulation 2016/679". WP 260 rev.01, 11 April 2018. ec.europa.eu/newsroom/article29/document.cfm?action=display&doc_id=51025 - endorsed by the EDPB

[25] The French Data Protection Authority has published several examples illustrating best practices in informing users as well as other transparency principles: https://design.cnil.fr/en/.

Moreover, necessary information should also be provided in the right context, at the appropriate time. Since the controller carries out many processing operations using the data collected on the website, a general privacy policy on the website alone is not sufficient for the controller to meet the requirements of transparency. The controller therefore designs an information flow, presenting the data subject with relevant information within the appropriate contexts using e.g. informational snippets or pop-ups. For example, when asking the data subject to enter personal data, the controller informs the data subject of how the personal data will be processed and why that personal data is necessary for the processing.

3.2 Lawfulness

67. The controller must identify a valid legal basis for the processing of personal data. Measures and safeguards should support the requirement to make sure that the whole processing lifecycle is in line with the relevant legal grounds of processing.

68. Key design and default elements for lawfulness may include:
- Relevance – The correct legal basis shall be applied to the processing.
- Differentiation[26] – The legal basis used for each processing activity shall be differentiated.
- Specified purpose – The appropriate legal basis must be clearly connected to the specific purpose of processing.[27]
- Necessity– Processing must be necessary and unconditional for the purpose to be lawful.
- Autonomy – The data subject should be granted the highest degree of autonomy as possible with respect to control over personal data within the frames of the legal basis.
- Gaining consent – consent must be freely given, specific, informed and unambiguous.[28] Particular consideration should be given to the capacity of children and young people to provide informed consent.
- Consent withdrawal – Where consent is the legal basis, the processing should facilitate withdrawal of consent. Withdrawal shall be as easy as giving consent. If not, then the consent mechanism of the controller does not comply with the GDPR.[29]
- Balancing of interests – Where legitimate interests is the legal basis, the controller must carry out a weighted balancing of interest, giving particular consideration to the power imbalance, specifically children under the age of 18 and other vulnerable groups. There shall be measures and safeguards to mitigate the negative impact on the data subjects.
- Predetermination – The legal basis shall be established before the processing takes place.
- Cessation – If the legal basis ceases to apply, the processing shall cease accordingly.
- Adjust – If there is a valid change of legal basis for the processing, the actual processing must be adjusted in accordance with the new legal basis.[30]

[26] EDPB. "Guidelines 2/2019 on the processing of personal data under Article 6(1)(b) GDPR in the context of the provision of online services to data subjects". Version 2.0, 8 October 2019. edpb.europa.eu/sites/edpb/files/files/file1/edpb_guidelines-art_6-1-b-adopted_after_public_consultation_en.pdf

[27] See section on purpose limitation below.

[28] See Guidelines 05/2020 on consent under Regulation 2016/679. https://edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-052020-consent-under-regulation-2016679_en

[29] See Guidelines 05/2020 on consent under Regulation 2016/679, p. 24. https://edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-052020-consent-under-regulation-2016679_en

[30] If the original legal basis is consent, see Guidelines 05/2020 on consent under Regulation 2016/679. https://edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-052020-consent-under-regulation-2016679_en

- Allocation of responsibility – Whenever joint controllership is envisaged, the parties must apportion in a clear and transparent way their respective responsibilities vis-à-vis the data subject, and design the measures of the processing in accordance with this allocation.

Example

A bank plans to offer a service to improve efficiency in the management of loan applications. The idea behind the service is that the bank, by requesting permission from the customer, is able to retrieve data about the customer directly from the public tax authorities. This example does not consider processing of personal data from other sources.

Obtaining personal data about the data subject's financial situation is necessary in order to take steps at the request of the data subject prior to entering into a loan contract.[31] However, gathering personal data directly from the tax administration is not considered necessary, because the customer is able to enter into a contract by providing the information from the tax administration him or herself. Although the bank may have a legitimate interest in acquiring the documentation from the tax authorities directly, for example to ensure efficiency in the loan processing, giving banks such direct access to the personal data of applicants presents a risks related to the use or potential misuse of access rights

When implementing the principle of lawfulness, the controller realizes that in this context, they cannot use the "necessary for contract" basis for the part of the processing that involves gathering personal data directly from the tax authorities. The fact that this specific processing presents a risk of the data subject becoming less involved in the processing of their data is also a relevant factor in assessing the lawfulness of the processing itself. The bank concludes that this part of the processing has to rely on another legal basis of processing. In the particular Member State where the controller is located, there are national laws that permits the bank to gather information from the public tax authorities directly, where the data subject consents to this beforehand.

The bank therefore presents information about the processing on the online application platform in such a manner that makes it easy for data subjects to understand what processing is mandatory and what is optional. The processing options, by default, do not allow retrieval of data directly from other sources than the data subject herself, and the option for direct information retrieval is presented in a manner that does not deter the data subject from abstaining. Any consent given to collect data directly from other controllers is a temporary right of access to a specific set of information.

Any given consent is processed electronically in a documentable manner, and data subjects are presented with an easy way of controlling what they have consented to and to withdraw their consent.

The controller has assessed these DPbDD requirements beforehand and includes all of these criteria in their requirements specification for the tender to procure the platform. The controller is aware that if they do not include the DPbDD requirements in the tender, it may either be too late or a very costly process to implement data protection afterwards.

3.3 Fairness

69. Fairness is an overarching principle which requires that personal data should not be processed in a way that is unjustifiably detrimental, unlawfully discriminatory, unexpected or misleading to the data

[31] See Article 6(1)(b) GDPR.

subject. Measures and safeguards implementing the principle of fairness also support the rights and freedoms of data subjects, specifically the right to information (transparency), the right to intervene (access, erasure, data portability, rectify) and the right to limit the processing (right not to be subject to automated individual decision-making and non-discrimination of data subjects in such processes).

70. Key design and default fairness elements may include:
    - Autonomy – Data subjects should be granted the highest degree of autonomy possible to determine the use made of their personal data, as well as over the scope and conditions of that use or processing.
    - Interaction – Data subjects must be able to communicate and exercise their rights in respect of the personal data processed by the controller.
    - Expectation – Processing should correspond with data subjects' reasonable expectations.
    - Non-discrimination – The controller shall not unfairly discriminate against data subjects.
    - Non-exploitation – The controller should not exploit the needs or vulnerabilities of data subjects.
    - Consumer choice – The controller should not "lock in" their users in an unfair manner. Whenever a service processing personal data is proprietary, it may create a lock-in to the service, which may not be fair, if it impairs the data subjects' possibility to exercise their right of data portability in accordance with Article 20.
    - Power balance – Power balance should be a key objective of the controller-data subject relationship. Power imbalances should be avoided. When this is not possible, they should be recognised and accounted for with suitable countermeasures.
    - No risk transfer – Controllers should not transfer the risks of the enterprise to the data subjects.
    - No deception – Data processing information and options should be provided in an objective and neutral way, avoiding any deceptive or manipulative language or design.
    - Respect rights – The controller must respect the fundamental rights of data subjects and implement appropriate measures and safeguards and not impinge on those rights unless expressly justified by law.
    - Ethical – The controller should see the processing's wider impact on individuals' rights and dignity.
    - Truthful – The controller must make available information about how they process personal data, they should act as they declare they will and not mislead the data subjects.
    - Human intervention – The controller must incorporate *qualified* human intervention that is capable of uncovering biases that machines may create in accordance with the right to not be subject to automated individual decision making in Article 22.[32]
    - Fair algorithms – Regularly assess whether algorithms are functioning in line with the purposes and adjust the algorithms to mitigate uncovered biases and ensure fairness in the processing. Data subjects should be informed about the functioning of the processing of personal data based on algorithms that analyse or make predictions about them, such as work performance, economic situation, health, personal preferences, reliability or behaviour, location or movements.[33]

[32] See Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679. https://ec.europa.eu/newsroom/article29/document.cfm?action=display&doc_id=49826
[33] See Recital 71 GDPR.

Example 1

A controller operates a search engine that processes mostly user-generated personal data. The controller benefits from having large amounts of personal data and being able to use that personal data for targeted advertisements. The controller therefore wishes to influence data subjects to allow more extensive collection and use of their personal data. Consent is to be collected by presenting processing options to the data subject.

When implementing the fairness principle, taking into account the nature, scope, context and purpose of the processing, the controller realizes that they cannot present the options in a way that nudges the data subject in the direction of allowing the controller to collect more personal data than if the options were presented in an equal and neutral way. This means that they cannot present the processing options in such a manner that makes it difficult for data subjects to abstain from sharing their data, or make it difficult for the data subjects to adjust their privacy settings and limit the processing. These are examples of dark patterns, which are contrary to the spirit of Article 25. The default options for the processing should not be invasive, and the choice for further processing should be presented in a manner that does not pressure the data subject to give consent. Therefore, the controller presents the options to consent or abstain as two equally visible choices, accurately representing the ramifications of each choice to the data subject.

Example 2

Another controller processes personal data for the provision of a streaming service where users may choose between a regular subscription of standard quality and a premium subscription with higher quality. As part of the premium subscription, subscribers get prioritized customer service.

With regard to the fairness principle, the prioritized customer service granted to premium subscribers cannot discriminate the regular subscribers' access to exercise their rights according to the GDPR Article 12. This means that although the premium subscribers get prioritized service, such prioritization cannot result in a lack of appropriate measures to respond to request from regular subscribers without undue delay and in any event within one month of receipt of the requests.

Prioritized customers may pay to get better service, but all data subjects shall have equal and indiscriminate access to enforce their rights and freedoms as required under Article 12.

3.4 Purpose Limitation[34]

71. The controller must collect data for specified, explicit, and legitimate purposes, and not further process the data in a manner that is incompatible with the purposes for which they were collected.[35] The design of the processing should therefore be shaped by what is necessary to achieve the purposes. If any

[34] The Article 29 Working Party provided guidance for the understanding of the principle of purpose limitation under Directive 95/46/EC. Although the Opinion is not adopted by the EDBP, it may still be relevant as the wording of the principle is the same under the GDPR. Article 29 Working Party. "Opinion 03/2013 on purpose limitation". WP 203, 2 April 2013. ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2013/wp203_en.pdf

[35] Art. 5(1)(b) GDPR.

further processing is to take place, the controller must first make sure that this processing has purposes compatible with the original ones and design such processing accordingly. Whether a new purpose is compatible or not, shall be assessed according to the criteria in Article 6(4).

72. Key design and default purpose limitation elements may include:

- Predetermination – The legitimate purposes shall be determined before the design of the processing.
- Specificity – The purposes shall be specified and explicit as to why personal data is being processed.
- Purpose orientation – The purpose of processing should guide the design of the processing and set processing boundaries.
- Necessity – The purpose determines what personal data is necessary for the processing.
- Compatibility – Any new purpose must be compatible with the original purpose for which the data was collected and guide relevant changes in design.
- Limit further processing – The controller should not connect datasets or perform any further processing for new incompatible purposes.
- Limitations of reuse – The controller should use technical measures, including hashing and encryption, to limit the possibility of repurposing personal data. The controller should also have organisational measures, such as policies and contractual obligations, which limit reuse of personal data.
- Review – The controller should regularly review whether the processing is necessary for the purposes for which the data was collected and test the design against purpose limitation.

Example

The controller processes personal data about its customers. The purpose of the processing is to fulfil a contract, i.e. to be able to deliver goods to the correct address and obtain payment. The personal data stored is the purchase history, name, address, e-mail address and telephone number.

The controller is considering buying a Customer Relationship Management (CRM) product that gathers all the customer data about sales, marketing and customer service in one place. The product gives the opportunity of storing all phone calls, activities, documents, emails and marketing campaigns to get a 360-degree view of the customer. Moreover, the CRM is capable of automatically analysing the customers' purchasing power by using public information. The purpose of the analysis is to better target advertising activities. Those activities do not form part of the original lawful purpose of the processing.

To be in line with the principle of purpose limitation, the controller requires the provider of the product to map the different processing activities that use personal data to the purposes relevant for the controller.

After receiving the results of the mapping, the controller assesses whether the new marketing purpose and the targeted advertisement purpose are compatible with the original purposes defined when the data was collected, and whether there is a sufficient legal basis for the respective processing. If the assessment does not return a positive answer, the controller shall not proceed to use the respective functionalities. Alternatively, the controller could choose to forego the assessment and simply not make use of the described functionalities of the product.

3.5 Data Minimisation

73. Only personal data that is adequate, relevant and limited to what is **necessary** for the purpose shall be processed.[36] As a result, the controller has to predetermine which features and parameters of processing systems and their supporting functions are permissible. Data minimisation substantiates and operationalises the principle of necessity. In the further processing, the controller should periodically consider whether processed personal data is still adequate, relevant and necessary, or if the data shall be deleted or anonymized.

74. Controllers should first of all determine whether they even need to process personal data for their relevant purposes. The controller should verify whether the relevant purposes can be achieved by processing less personal data, or having less detailed or aggregated personal data or without having to process personal data at all[37]. Such verification should take place before any processing takes place, but could also be carried out at any point during the processing lifecycle. This is also consistent with Article 11.

75. Minimising can also refer to the degree of identification. If the purpose of the processing does not require the final set of data to refer to an identified or identifiable individual (such as in statistics), but the initial processing does (e.g. before data aggregation), then the controller shall delete or anonymize personal data as soon as identification is no longer needed. Or, if continued identification is needed for other processing activities, personal data should be pseudonymized to mitigate risks for the data subjects' rights.

76. Key design and default data minimisation elements may include:
    - Data avoidance – Avoid processing personal data altogether when this is possible for the relevant purpose.
    - Limitation – Limit the amount of personal data collected to what is necessary for the purpose
    - Access limitation – Shape the data processing in a way that a minimal number of people need access to personal data to perform their duties, and limit access accordingly.
    - Relevance – Personal data should be relevant to the processing in question, and the controller should be able to demonstrate this relevance.
    - Necessity – Each personal data category shall be necessary for the specified purposes and should only be processed if it is not possible to fulfil the purpose by other means.
    - Aggregation – Use aggregated data when possible.
    - Pseudonymization – Pseudonymize personal data as soon as it is no longer necessary to have directly identifiable personal data, and store identification keys separately.
    - Anonymization and deletion – Where personal data is not, or no longer necessary for the purpose, personal data shall be anonymized or deleted.
    - Data flow – The data flow should be made efficient enough to not create more copies than necessary.
    - "State of the art" – The controller should apply up to date and appropriate technologies for data avoidance and minimisation.

[36] Art. 5(1)(c) GDPR.
[37] Recital 39 GDPR so states: "...Personal data should be processed only if the purpose of the processing could not reasonably be fulfilled by other means."

Example 1

A bookshop wants to add to their revenue by selling their books online. The bookshop owner wants to set up a standardised form for the ordering process. To ensure customers fill out all the wanted information the bookshop owner makes all of the fields in the form mandatory (if you don't fill out all the fields the customer can't place the order). The webshop owner initially uses a standard contact form, which asks information including the customer's date of birth, phone number and home address. However, not all the fields in the form are necessary for the purpose of buying and delivering the books. In this particular case, if the data subject pays for the product up front, the data subject's date of birth and phone number are not necessary for the purchase of the product. This means that these cannot be required fields in the web form to order the product, unless the controller can clearly demonstrate that it is otherwise necessary, and why the fields are necessary. Moreover, there are situations where an address will not be necessary. For example, when ordering an eBook the customer can download the product directly to their device.

The webshop owner therefore decides to make two web forms: one for ordering books, with a field for the customer's address and one web form for ordering eBooks without a field for the customer's address.

Example 2

A public transportation company wishes to gather statistical information based on travellers' routes. This is useful for the purposes of making proper choices on changes in public transport schedules and proper routings of the trains. The passengers have to pass their ticket through a reader every time they enter or exit a means of transport. Having carried out a risk assessment related to the rights and freedoms of passengers' regarding the collection of passengers' travel routes, the controller establishes that it is possible to identify the passengers in circumstances where they live or work in scarcely populated areas, based on single route identification thanks to the ticket identifier. Therefore, since it is not necessary for the purpose of optimizing the public transport schedules and routings of the trains, the controller does not store the ticket identifier. Once the trip is over, the controller only stores the individual travel routes so as to not be able to identify trips connected to a single ticket, but only retains information about separate travel routes.

In cases where there can still be a risk of identifying a person solely by their public transportation travel route the controller implements statistical measures to reduce the risk, such as cutting the beginning and the end of the route.

Example 3

A courier aims at assessing the effectiveness of its deliveries in terms of delivery times, workload scheduling and fuel consumption. In order to reach this goal, the courier has to process a number of personal data relating to both employees (drivers) and customers (addresses, items to be delivered, etc.). This processing operation entails risks of both monitoring employees, which requires specific legal safeguards, and tracking customers' habits through the knowledge of the delivered items over time. These risks can be significantly reduced with appropriate pseudonymization of employees and customers. In particular if pseudonymization keys are frequently rotated and macro areas are considered instead of detailed addresses, an effective data minimisation is pursued, and the controller

can solely focus on the delivery process and on the purpose of resource optimization, without crossing the threshold of monitoring individuals' (customers' or employees') behaviours.

Example 4

A hospital is collecting data about its patients in a hospital information system (electronic health record). Hospital staff needs to access patient files to inform their decisions regarding care for and treatment of the patients, and for the documentation of all diagnostic, care and treatment actions taken. By default, access is granted to only those members of the medical staff who are assigned to the treatment of the respective patient in the speciality department she or he is assigned to. The group of people with access to a patient's file is enlarged if other departments or diagnostic units are involved in the treatment. After the patient is discharged, and billing is completed, access is reduced to a small group of employees per speciality department who answer requests for medical information or a consultation made or asked for by other medical service providers upon authorization by the respective patient.

3.6 Accuracy

77. Personal data shall be accurate and kept up to date, and every reasonable step shall be taken to ensure that personal data that is inaccurate, having regard to the purposes for which they are processed, are erased or rectified without delay.[38]

78. The requirements should be seen in relation to the risks and consequences of the concrete use of data. Inaccurate personal data could be a risk to the data subjects' rights and freedoms, for example when leading to a faulty diagnosis or wrongful treatment of a health protocol, or an incorrect image of a person can lead to decisions being made on the wrong basis either manually, using automated decision-making, or through artificial intelligence.

79. Key design and default accuracy elements may include:
- Data source – Sources of personal data should be reliable in terms of data accuracy.
- Degree of accuracy – Each personal data element should be as accurate as necessary for the specified purposes.
- Measurably accurate - Reduce the number of false positives/negatives, for example biases in automated decisions and artificial intelligence.
- Verification – Depending on the nature of the data, in relation to how often it may change, the controller should verify the correctness of personal data with the data subject before and at different stages of the processing (e.g. to age requirements).
- Erasure/rectification – The controller shall erase or rectify inaccurate data without delay. The controller shall in particular facilitate this where the data subjects are or were children and later want to remove such personal data.[39]
- Error propagation avoidance – Controllers should mitigate the effect of an accumulated error in the processing chain.

[38] Art. 5(1)(d) GDPR.
[39] Cf. Recital 65.

- Access – Data subjects should be given information about and effective access to personal data in accordance with the GDPR articles 12 to 15 in order to control accuracy and rectify as needed.
- Continued accuracy – Personal data should be accurate at all stages of the processing, tests of accuracy should be carried out at critical steps.
- Up to date – Personal data shall be updated if necessary for the purpose.
- Data design - Use of technological and organisational design features to decrease inaccuracy, for example present concise predetermined choices instead of free text fields.

Example 1

An insurance company wishes to use artificial intelligence (AI) to profile customers buying insurance as a basis for their decision making when calculating the insurance risk. When determining how their AI solutions should be developed, they are determining the means of processing and shall consider data protection by design when choosing an AI application from a vendor and when deciding on how to train the AI.

When determining how to train the AI, the controller should have accurate data to achieve precise results. Therefore, the controller should ensure that the data used to train the AI is accurate.

Granted that they have a valid legal basis to train the AI using personal data from a large subset of their existing customers, the controller chooses a pool of customers that is representative of the population to also avoid bias.

The customer data is then collected from the respective data handling system, including data on the type of insurance, for example health insurance, home insurance, travel insurance, etc. as well as data from public registries they have lawful access to. All data are pseudonymized prior to transfer to the system dedicated to the training of the AI model.

To ensure that the data used for AI training is as accurate as possible, the controller only collects data from data sources with correct and up-to date information.

The insurance company tests whether the AI is reliable and provides non-discriminatory results both during its development and finally before the product is released. When the AI is fully trained and operative, the insurance company uses the results to support the insurance risk assessments, yet without solely relying on the AI to decide whether to grant insurance, unless the decision is made in accordance with the exceptions in Article 22 (2) GDPR.

The insurance company will also regularly review the results from the AI, to maintain the reliability and when necessary adjust the algorithm.

Example 2

The controller is a health institution looking to find methods to ensure the integrity and accuracy of personal data in their client registers.

In situations where two persons arrive at the institution at the same time and receive the same treatment, there is a risk of mistaking them if the only parameter to distinguish them is by name. To ensure accuracy, the controller needs a unique identifier for each person, and therefore more information than just the name of the client.

The institution uses several systems containing personal information of clients, and needs to ensure that the information related to the client is correct, accurate and consistent in all the systems at any point in time. The institution has identified several risks that may arise if information is changed in one system but not in the others.

The controller decides to mitigate the risk by using a hashing technique that can be used to ensure integrity of data in the treatment journal. Immutable cryptographic time stamps are created for treatment journal records and the client associated with them so that any changes can be recognized, correlated and traced if required.

3.7 Storage limitation

80. The controller must ensure that personal data is kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data is processed.[40] It is vital that the controller knows exactly what personal data the company processes and why. The purpose of the processing shall be the main criterion to decide in how long personal data shall be stored.

81. Measures and safeguards that implement the principle of storage limitation shall complement the rights and freedoms of the data subjects, specifically, the right to erasure and the right to object.

82. Key design and default storage limitation elements may include:
    - Deletion and anonymization – The controller should have clear internal procedures and functionalities for deletion and/or anonymization.
    - Effectiveness of anonymization/deletion – The controller shall make sure that it is not possible to re-identify anonymized data or recover deleted data, and should test whether this is possible.
    - Automation – Deletion of certain personal data should be automated
    - Storage criteria – The controller shall determine what data and length of storage is necessary for the purpose.
    - Justification – The controller shall be able to justify why the period of storage is necessary for the purpose and the personal data in question, and be able to disclose the rationale behind, and legal grounds for the retention period.
    - Enforcement of retention policies – The controller should enforce internal retention policies and conduct tests of whether the organization practices its policies.
    - Backups/logs – Controllers shall determine what personal data and length of storage is necessary for back-ups and logs.
    - Data flow – Controllers should beware of the flow of personal data, and the storage of any copies thereof, and seek to limit their "temporary" storage.

Example

The controller collects personal data where the purpose of the processing is to administer a membership of the data subject. The personal data shall be deleted when the membership is terminated and there is no legal basis for further storage of the data.

[40] Art. 5(1)(c) GDPR.

The controller first draws up an internal procedure for data retention and deletion. According to this, employees shall manually delete personal data after the retention period ends. The employee follows the procedure to regularly delete and correct data from any devices, from backups, logs, e-mails and other relevant storage media.

To make deletion more effective, and less error-prone, the controller then implements an automatic system instead, in order to delete data automatically, reliably and more regularly. The system is configured to follow the given procedure for data deletion which then occurs at a predefined regular interval to remove personal data from all of the company's storage media. The controller reviews and tests the retention procedure regularly and ensures that it concurs with the up-to-date retention policy.

3.8 Integrity and confidentiality

83. The principle of integrity and confidentiality includes protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures. The security of personal data requires appropriate measures designed to prevent and manage data breach incidents; to guarantee the proper execution of data processing tasks, and compliance with the other principles; and to facilitate the effective exercise of individuals' rights.

84. Recital 78 states that one of the DPbDD measures could consist of enabling the controller to *"create and improve security features"*. Along with other DPbDD measures, Recital 78 suggests a responsibility on the controllers to continually assess whether it is using the appropriate means of processing at all times and to assess whether the chosen measures actually counter the existing vulnerabilities. Furthermore, controllers should conduct regular reviews of the information security measures that surround and protect personal data, and the procedure for handling data breaches.

85. Key design and default integrity and confidentiality elements may include:
    - Information security management system (ISMS) – Have an operative means of managing policies and procedures for information security.
    - Risk analysis – Assess the risks against the security of personal data by considering the impact on individuals' rights and counter identified risks. For use in risk assessment; develop and maintain a comprehensive, systematic and realistic "threat modelling" and an attack surface analysis of the designed software to reduce attack vectors and opportunities to exploit weak points and vulnerabilities.
    - Security by design – Consider security requirements as early as possible in the system design and development and continuously integrate and perform relevant tests.
    - Maintenance – Regular review and test software, hardware, systems and services, etc. to uncover vulnerabilities of the systems supporting the processing.
    - Access control management – Only the authorized personnel who need to should have access to the personal data necessary for their processing tasks, and the controller should differentiate between access privileges of authorized personnel.
        - Access limitation (agents) – Shape the data processing in a way that a minimal number of people need access to personal data to perform their duties, and limit access accordingly.
        - Access limitation (content) – In the context of each processing operation, limit access to only those attributes per data set that are needed to perform that operation.

Moreover, limit access to data pertaining to those data subjects who are in the remit of the respective employee.

- Access segregation – Shape the data processing in a way that no individual needs comprehensive access to all data collected about a data subject, much less all personal data of a particular category of data subjects.

- Secure transfers – Transfers shall be secured against unauthorized and accidental access and changes.
- Secure storage – Data storage shall be secure from unauthorized access and changes. There should be procedures to assess the risk of centralized or decentralized storage, and what categories of personal data this applies to. Some data may need additional security measures than others or isolation from others.
- Pseudonymization – Personal data and back-ups/logs should be pseudonymized as a security measure to minimise risks of potential data breaches, for example using hashing or encryption.
- Backups/logs – Keep back-ups and logs to the extent necessary for information security, use audit trails and event monitoring as a routine security control. These shall be protected from unauthorised and accidental access and change and reviewed regularly and incidents should be handled promptly.
- Disaster recovery/ business continuity – Address information system disaster recovery and business continuity requirements to restore the availability of personal data following up major incidents.
- Protection according to risk – All categories of personal data should be protected with measures adequate with respect to the risk of a security breach. Data presenting special risks should, when possible, be kept separated from the rest of the personal data.
- Security incident response management – Have in place routines, procedures and resources to detect, contain, handle, report and learn from data breaches.
- Incident management – Controller should have processes in place to handle breaches and incidents, in order to make the processing system more robust. This includes notification procedures, such as management of notification (to the supervisory authority) and information (to data subjects).

Example

A controller wants to extract large quantities of personal data from a medical database containing electronic (patient) health records to a dedicated database server in the company in order to process the extracted data for quality assurance purposes. The company has assessed the risk for routing the extracts to a server that is accessible to all of the company's employees as likely to be high for data subjects' rights and freedoms. Since there is only one department in the company who needs to process the patient data extracts, the controller decides to restrict access to the dedicated server to employees in that department. Moreover, to further reduce risk, the data will be pseudonymized before they are transferred.

To regulate access and mitigate possible damage from malware, the company decides to segregate the network, and establish access controls to the server. In addition, they put up security monitoring and an intrusion detection and prevention system and isolates it from routine use. An automated auditing system is put in place to monitor access and changes. Reporting and automated alerts are generated from this when certain events related to usage are configured. The controller will ensure that users only have access on a need to know basis and with the appropriate access level. Inappropriate use can be quickly and easily detected.

Some of the extracts have to be compared with new extracts, and therefore are required to be stored for three months. The controller decides to put them into separate databases on the same server, and use both transparent and column-level encryption to store them. Keys for column data decryption are stored in dedicated security modules that can only be used by authorized personnel, but not extracted.

Handling upcoming incidents makes the system more robust, and reliable. The data controller understands that preventative and effective measures and safeguards should be built into all personal data processing it undertakes now and in the future, and that doing so may help prevent future such data breach incidents.

The controller establishes these security measures both to ensure accuracy, integrity and confidentiality, but also to prevent malware spread by cyber-attacks and to make the solution robust. Having robust security measures contributes to build trust with the data subjects.

3.9 Accountability[41]

86. The principle of accountability states that the controller shall be responsible for, and be able to demonstrate compliance with all of the abovementioned principles.

87. The controller needs to be able to demonstrate compliance with the principles. In doing so, the controller may demonstrate the effects of the measures taken to protect the data subjects' rights, and why the measures are considered to be appropriate and effective. For example, demonstrating why a measure is appropriate to ensure the principle of storage limitation in an effective manner.

88. To be able to process personal data responsibly, the controller should have both the knowledge of and the ability to implement data protection. This entails that the controller should understand their data protection obligations of the GDPR and be able to comply with these obligations.

4 ARTICLE 25(3) CERTIFICATION

89. According to Article 25(3), certification pursuant to Article 42 may be used as an element to demonstrate compliance with DPbDD. Conversely, documents demonstrating compliance with DPbDD may also be useful in a certification process. This means that where a processing operation by a controller or a processor has been certified as per Article 42, supervisory authorities shall take this into account in their assessment of compliance with the GDPR, specifically with regards to DPbDD.

90. When a processing operation by a controller or processor is certified according to Article 42, the elements that contribute to demonstrating compliance with Article 25(1) and (2) are the design processes, i.e. the process of determining the means of processing, the governance and the technical and organizational measures to implement the data protection principles The data protection certification criteria are determined by the certification bodies or certification scheme owners and then approved by the competent supervisory authority or by the EDPB. For further information about certification mechanisms, we refer the reader to the EDPB Guideline on Certification[42] and other relevant guidance, as published on the EDPB website.

[41] See Recital 74, where controllers are required to demonstrate the effectiveness of their measures.
[42] EDPB. "Guidelines 1/2018 on certification and identifying certification criteria in accordance with Articles 42 and 43 of the Regulation". Version 3.0, 4 June 2019.
edpb.europa.eu/sites/edpb/files/files/file1/edpb_guidelines_201801_v3.0_certificationcriteria_annex2_en.pdf

91. Even where a processing operation is awarded a certification in accordance with Article 42, the controller still has the responsibility to continuously monitor and improve compliance with the DPbDD-criteria of Article 25.

5 ENFORCEMENT OF ARTICLE 25 AND CONSEQUENCES

92. Supervisory authorities may assess compliance with Article 25 according to the procedures listed in Article 58. The corrective powers are specified in Article 58(2) and include the issuance of warnings, reprimands, orders to comply with data subjects' rights, limitations on or ban of processing, administrative fines, etc.

93. DPbDD is further a factor in determining the level of monetary sanctions for breaches of the GDPR, see Article 83(4).[43] [44]

6 RECOMMENDATIONS

94. Although not directly addressed in Article 25, processors and producers are also recognized as key enablers for DPbDD, they should be aware that controllers are required to only process personal data with systems and technologies that have built-in data protection.

95. When processing on behalf of controllers, or providing solutions to controllers, processors and producers should use their expertise to build trust and guide their customers, including SMEs, in designing /procuring solutions that embed data protection into the processing. This means in turn that the design of products and services should facilitate controllers' needs.

96. It should be kept in mind when implementing Article 25 that the main design objective is the *effective implementation* of the principles and *protection* of the rights of data subjects into the appropriate measures of the processing. In order to facilitate and enhance the adoption of DPbDD, we make the following recommendations to controllers as well as producers and processors:

- Controllers should think of data protection from the *initial stages* of planning a processing operation, even before the time of determination of the means of processing.
- Where the controller has a Data Protection Officer (DPO), the EDPB encourages the active involvement of the DPO to integrate DPbDD in the procurement and development procedures, as well as in the whole processing life-cycle.
- A processing operation may be *certified*. The ability to get a processing operation certified provides an added value to a controller when choosing between different processing software, hardware, services and/or systems from producers or processors. Therefore, producers should strive to demonstrate DPbDD in the life-cycle of their development of a processing solution. A certification seal may also guide data subjects in their choice between different goods and services. Having the ability to get a processing certified can serve as a competitive advantage for producers, processors and controllers, and even enhances data subjects' trust in the

[43] Article 83(2)(d) GDPR stipulates that in determining the imposition of fines for breach of the GDPR *"due regard" shall be taken of "the degree of responsibility of the controller or processor taking into account technical and organisational measures implemented by them pursuant to Articles 25 and 32".*

[44] More information on fines can be found in Article 29 Working Party. "Guidelines on application and setting of administrative fines for the purposes of the Regulation 2016/679". WP 253, 3 October 2017. ec.europa.eu/newsroom/just/document.cfm?doc_id=47889 - endorsed by the EDPB.

processing of their personal data. If no certification is offered, controllers should seek to have other *guarantees* that producers or processors comply with the requirements of DPbDD.

- Controllers, processors and producers, should consider their obligations to provide children under 18 and other vulnerable groups with specific protection in complying with DPbDD.
- Producers and processors should seek to facilitate DPbDD implementation in order to support the controller's ability to comply with Article 25 obligations. Controllers, on the other hand, should not choose producers or processors who do not offer systems enabling or supporting the controller to comply with Article 25, because controllers will be held accountable for the lack of implementation thereof.
- Producers and processors should play an active role in ensuring that the criteria for the "state of the art" are met, and notify controllers of any changes to the "state of the art" that may affect the effectiveness of the measures they have in place. Controllers should include this requirement as a contractual clause to make sure they are kept up to date.
- The EDPB recommends controllers to require that producers and processors demonstrate how their hardware, software, services or systems enable the controller to comply with the requirements to accountability in accordance with DPbDD, for example by using key performance indicators to demonstrate the effectiveness of the measures and safeguards at implementing the principles and rights.
- The EDPB emphasizes the need for a harmonized approach to implement principles and rights in an effective manner and encourages associations or bodies preparing codes of conduct in accordance with Article 40 to also incorporate sector-specific guidance on DPbDD.
- Controllers should be fair to data subjects and transparent on how they assess and demonstrate effective DPbDD implementation, in the same manner as controllers demonstrate compliance with the GDPR under the principle of accountability.
- Privacy-enhancing technologies (PETs) that have reached the state-of-the-art maturity can be employed as a measure in accordance with the DPbDD requirements if appropriate in a risk based approach. PETs in themselves do not necessarily cover the obligations of Article 25. Controllers shall assess whether the measure is appropriate and effective in implementing the data protection principles and the rights of data subjects.
- Existing legacy systems are under the same DPbDD-obligations as new systems. If legacy systems do not already comply with DPbDD, and changes cannot be made to comply with the obligations, then the legacy system simply does not meet GDPR-obligations and cannot be used to process personal data.
- Article 25 does not lower the threshold of requirements for SMEs. The following points may facilitate SMEs' compliance with Article 25:
    - Do early risk assessments
    - Start with small processing – then scale its scope and sophistication later
    - Look for producer and processor guarantees of DPbDD, such as certification and adherence to code of conducts
    - Use partners with a good track record
    - Talk with DPAs
    - Read guidance from DPAs and the EDPB
    - Adhere to codes of conduct where available
    - Get professional help and advice

For the European Data Protection Board

The Chair

(Andrea Jelinek)